This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

Rank-one Solutions for SDP Relaxation of QCQPs in Power Systems

Tian Liu, *Student Member, IEEE,* Bo Sun, *Student Member, IEEE,* and Danny H.K. Tsang, *Fellow, IEEE*

***Abstract*—It has been shown that a large number of computationally difficult problems can be equivalently reformulated into quadratically constrained quadratic programs (QCQPs) in the literature of power systems. Due to the NP-hardness of general QCQPs, main effort of this stream of problems has been put into deriving near-optimal solutions with low computational complexity. Recently, semidefinite programming (SDP) relaxation has been recognized as a promising technique to solve QCQPs from various applications such as the alternating current (AC) optimal power flow (OPF) problem. However, this technique has not been guaranteed to achieve a rank-one solution, which is a necessary condition to recover a feasible solution of the original QCQPs. In this paper, instead of investigating the conditions under which a rank-one solution exists, we propose a general solution framework to derive near-optimal but rank-one solutions for the SDP relaxation of QCQPs. In the proposed algorithm, all the parameters are provided in a systematic manner. In order to demonstrate the effectiveness of our method, the proposed algorithm is applied to solve the AC-OPF and state estimation problems in various settings. Extensive numerical results show that our method succeeds in obtaining rank-one solutions in all our case studies and only small optimality gaps are induced by our approach.**

***Index Terms*—Quadratically constrained quadratic program (QCQP), semidefinite programming (SDP), rank approximation.**

I. Introduction

MANY problems in power systems can be formulated as quadratically constrained quadratic programs (QCQPs). One example is the alternating current (AC) optimal power flow (OPF) problem [1]–[3], which is the core problem of power system operation and control. It can be formulated as a QCQP since the power flow equations are quadratic in bus voltages when the rectangular form of the complex voltage is used [1]. Another example is the strategic bidding problem in the electricity market with a bi-level optimization formulation [4], [5]. This bi-level problem can be equivalently transformed to a single-level QCQP by replacing its lower-level problem, which is a linear program, with its Karush-Kuhn-Tucker (KKT) optimality conditions. In general, QCQPs are known to be NP-hard [6] and computationally difficult. Thus, a main stream of literature has been focusing on developing low-complexity algorithms to achieve near-optimal solutions instead of searching for the optimal solutions.

Recently, the semidefinite programming (SDP) relaxation technique has been used widely to solve QCQPs [7]–[9]. In [7], the authors propose a spatial branch-and-cut (SBC) approach to solve QCQPs with complex bounded variables (CQCQP). They use a complex SDP formulation strengthened with valid inequalities that are derived from the convex hull description of a nonconvex set of 2×2 positive semidefinite Hermitian matrices subject to a rank-one constraint. Their proposed branching methods based on the rank-one constraint can result in better performance compared to the benchmark reliability branching method. In [8], a systematic study of the quadratic convex (QC) relaxation for AC optimal power flow problems is presented. Their main theoretical result shows that the QC relaxation is stronger than the second-order cone (SOC) relaxation and neither dominates nor is dominated by the SDP relaxation. In [9], the authors introduce an exact reformulation of the SDP relaxation for OPF problems in order to deal with the issue that SDP solvers usually suffer from a lack of scalability. The proposed formulation is constrained by a set of polynomial constraints defined in the space of real variables. The new constraints can be seen as "cuts", strengthening weaker second-order cone relaxations and they show that a significant gain in computational efficiency can be achieved by their method compared to a standard SDP solver approach.

By solving the SDP relaxation problem of a QCQP, a lower bound of the original minimization problem can be achieved. Furthermore, it has been shown that the gap between the lower bound and the optimal objective value is bounded under specific conditions [10]–[12].

However, one important issue of the SDP relaxation technique is how to recover a feasible solution of the original QCQP from the solution of the SDP relaxation problem. Particularly, if the optimal solution of the SDP relaxation problem is rank-one, the optimal solution of the original QCQP can be reconstructed accordingly by a standard method [12]. Nevertheless, it is not guaranteed that the SDP relaxation problem can achieve rank-one solutions. In the example of OPF-based optimization problems [13], [14], the sufficient conditions for the existence of rank-one solutions by SDP relaxation have been studied. However, these conditions cannot always be met, especially when there are strict line-flow limits [15]. In that case, how to obtain a rank-one solution with a small loss of optimality is an inevitable yet challenging problem. An intuitive approach is to project the high rank solution to the nearest rank-one matrix, but the quality of the resulting solution is not guaranteed. Another technique is to enforce the SDP relaxation problem to achieve low-rank solutions by incorporating a rank penalty term into its objective [16]. In the literature, the rank penalty functions are chosen as the total reactive power generation [16], the apparent power loss [2] or the functions derived from the

T. Liu, B. Sun and D.H.K. Tsang are with the Hong Kong University of Science and Technology, Hong Kong (e-mail: {tliuai, bsunaa, eetsang}@ust.hk).

network Laplacian matrix [17]. However, these methods are just heuristics to enforce low rank solutions, and thus, they are not guaranteed to yield rank-one solutions. Molzahn *et al.* [18] apply a high-order moment relaxation method, which is a generalization of the SDP relaxation, to solve OPF problems. But feasible solutions can only be obtained in the cases with sufficiently high relaxation orders, which are computational challenging. You *et al.* [19] try to obtain a rank-one solution of the OPF problems in a heuristic way based on alternating direction method of multipliers (ADMM). In each iteration, they propose to solve a convex program without the rank constraint, followed by a polishing step to enforce rank-one solutions through singular value decomposition. However, their method cannot be guaranteed to converge because of the non-convexity of OPF problems.

In this paper, a solution framework is proposed to obtain rank-one solutions for the SDP relaxation of QCQPs. Our basic idea is also to add a rank penalty to the objective of the SDP relaxation problem to enforce a low rank solution. Instead of using the heuristic penalty functions in the existing works [2], [16], [17], we propose to approximate the rank penalty with some smooth functions, which are controllable in accuracy, and apply the majorization-miminization (MM) technique to iteratively solve the approximate problem. Compared with those penalization methods for OPF-based problems in the literature, the main advantages of our approach can be summarized as follows.

1) Our framework is generic and not limited to OPF problems because we directly approximate the rank function, which is not problem specific at all. In fact, any problem that can be equivalently formulated into QCQPs can be solved by the proposed method.
2) Both feasibility and convergence are guaranteed by our method. In particular, the optimal solution of the SDP relaxation with the approximate rank penalty converges to that of the original problem as the approximation parameter goes to zero. Moreover, it is ensured that our proposed algorithm converges to a feasible rank-one solution, which is at least a stationary point of the original problem.
3) The penalization parameters are provided in a systematic manner, which facilitates the use of this method, in contrast to those methods requiring the choice of appropriate penalization parameters [17].
4) A global exploration method is proposed to improve the solution quality by a local smoothing technique.

The rest of the paper is organized as follows. In Section II, the SDP relaxation for QCQPs is provided. The proposed rank-one solution framework based on rank penalty approximation and the MM technique are provided in details in Section III. The global search based on the local smoothing technique is introduced in Section IV. Case studies with application to OPF problems are presented in Section V and the application of our method on the state estimation problems is given in Section VI, followed by conclusions in Section VII.

Notations: $\mathbb{R}^n$, $\mathbb{R}^n_+$, $\mathbb{S}^n$ and $\mathbb{S}^n_+$ denote the n-dimensional real space, the set of n-dimensional non-negative real vectors, the set of real symmetric $n \times n$ matrices and the set of real $n \times n$ positive semidefinite (PSD) matrices, respectively. Superscripts T and $*$ denote the transpose operator and conjugate transpose operator, respectively. Tr denotes the matrix trace operator and the inner product on matrix space is denoted by $\langle \boldsymbol{W}, \boldsymbol{V} \rangle = \text{Tr}(\boldsymbol{W}^T \boldsymbol{V})$. $\text{Re}\{x\}$ and $\text{Im}\{x\}$ denote the real and imaginary part of x, respectively. The Frobenius norm and the nuclear norm are denoted by $||\boldsymbol{W}||_F = \sqrt{\langle \boldsymbol{W}, \boldsymbol{W} \rangle}$ and $||\boldsymbol{W}||_* = \text{Tr}(\sqrt{\boldsymbol{W}^* \boldsymbol{W}})$, respectively. The standard basis vectors in $\mathbb{R}^n$ are denoted as $e_1, e_2, \dots e_n$ and $\text{diag}(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^n$ is a diagonal matrix whose diagonal entries are the elements of $\boldsymbol{x}$.

II. SDP RELAXATION FOR QCQPS

A QCQP is an optimization problem whose objective function and constraints are both quadratic functions. For simplicity, we focus on the SDP relaxation for homogeneous QCQPs since all the inhomogeneous problems can be homogenized (as shown in Appendix A). The standard form of homogeneous QCQPs is as follows:

$$\begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} & \boldsymbol{x}^T \boldsymbol{A}_0 \boldsymbol{x} \\ \text{subject to} & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \le b_i,\ i = 1, \dots, m, \end{array} \tag{1}$$

where $b_i \in \mathbb{R}$ and $A_0, A_i \in \mathbb{S}^n$. In general, QCQPs are nonconvex and they are convex only when all the $\boldsymbol{A}_i \succeq 0, i = 0, \dots, m$, which means they are PSD matrices.

By noting that

$$\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} = \text{Tr}(\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x}) = \text{Tr}(\boldsymbol{A}_i \boldsymbol{x} \boldsymbol{x}^T),$$

we introduce a new matrix variable $\boldsymbol{W} = \boldsymbol{x}\boldsymbol{x}^T$ and Problem (1) can be equivalently formulated in terms of $\boldsymbol{W}$:

$$\begin{array}{ll} \underset{\boldsymbol{W} \in \mathbb{S}^n}{\text{minimize}} & \text{Tr}(\boldsymbol{A}_0 \boldsymbol{W}) \\ \text{subject to} & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i \boldsymbol{W}) \le b_i,\ i = 1, \dots, m, \\ & \text{rank}(\boldsymbol{W}) = 1. \end{array} \tag{2}$$

By removing the nonconvex rank constraint, Problem (2) is relaxed into an SDP, namely,

$$\begin{array}{ll} \underset{\boldsymbol{W} \in \mathbb{S}^n}{\text{minimize}} & \text{Tr}(\boldsymbol{A}_0 \boldsymbol{W}) \\ \text{subject to} & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i \boldsymbol{W}) \le b_i,\ i = 1, \dots, m. \end{array} \tag{3}$$

Note that being the SDP relaxation of Problem (2), Problem (3) is convex and can be solved efficiently by existing commercial solvers. In general, the optimal objective value of the SDP relaxation (3) is a lower bound of the optimal value of the original Problem (2). If the optimal solution of the SDP relaxation, $\boldsymbol{W}^{opt}$, is rank-one, the optimal solution of the original QCQP, $\boldsymbol{x}^{opt}$, can be recovered from $\boldsymbol{W}^{opt}$ by solving $\boldsymbol{W}^{opt} = \boldsymbol{x}^{opt} (\boldsymbol{x}^{opt})^T$. However, if the rank of $\boldsymbol{W}^{opt}$ is larger than one, it is a critical issue to convert $\boldsymbol{W}^{opt}$ into the optimal solution of the original QCQP. In fact, it is even nontrivial to transform $\boldsymbol{W}^{opt}$ into a feasible solution of the QCQP. In the following section, we will introduce our rank penalization method to induce a rank-one near-optimal solution of Problem (2).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

III. Rank-one Solution Framework for SDP Relaxation

In this section, we first introduce the approximation approach for the rank function and then, leverage the MM method to solve the approximate problem.

A. Approximating the rank function

In order to enforce a low rank solution of the SDP relaxation, we add the rank of the matrix variable into the objective function with a weighting factor $\eta > 0$ and the resulting problem is as follows:

$$\begin{aligned} \underset{\boldsymbol{W}\in\mathbb{S}^n}{\text{minimize}} \quad & \text{Tr}(\boldsymbol{A}_0\boldsymbol{W}) + \eta \cdot \text{rank}(\boldsymbol{W}) \\ \text{subject to} \quad & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i\boldsymbol{W}) \le b_i,\ i = 1, \ldots, m. \end{aligned} \tag{4}$$

By defining Problem (4), we move the difficulty brought by the nonconvex rank constraint to the objective. Note that this approach can be considered as a Lagrangian relaxation of the rank-one constraint if we take η as the dual variable of the rank-one constraint [20], [21]. This can be clearer if we realize the implicit fact that rank($\boldsymbol{W}$) ≥ 1, and we only need to enforce the constraint rank($\boldsymbol{W}$) ≤ 1. However, rather than relying on any strong duality condition, we are considering our approach as the *penalty method* [22], where we just require η to be sufficiently large, but not necessarily to be the dual variable. Moreover, it can be easily shown that for a large enough η, the optimal solution of Problem (4) is the same as that of Problem (2) under very mild conditions: i) Problem (3) is bounded; ii) there exists a rank-one feasible point and iii) η is sufficiently large. However, Problem (4) is still hard to solve due to the fact that the rank function is non-differentiable and discontinuous, which will be shown later. Therefore, we propose to approximate the rank function by a summation of parameterized concave functions and show that the optimal solution of the approximate problem converges to the optimal solution of Problem (4) as the approximation parameter approaches zero [23], [24].

Note that for any $\boldsymbol{W} \in \mathbb{S}^n$, rank($\boldsymbol{W}$) $= \sum_{i=1}^n u(\sigma_i(\boldsymbol{W}))$, where $\sigma_i(\boldsymbol{W})$ denotes the ith largest singular value of $\boldsymbol{W}$ and $u(x)$ denotes the unit step function for $x \ge 0$ defined by $u(x) = 1$ if $x > 0$ and $u(x) = 0$ if $x = 0$. $u(x)$ is discontinuous, non-differentiable, and thus difficult to minimize directly. In order to handle this, we approximate it by a continuous and differentiable function $\theta(x, \varepsilon)$, in which the parameter $\varepsilon > 0$ is used to control the accuracy of this approximation. There are many choices of the approximate function, among which we choose $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$, because it is not only differentiable but also a lower bound of $u(x)$ and $\lim_{\varepsilon \to 0} \theta(x, \varepsilon) = u(x)$. These good properties facilitate the design of our solution framework. In Fig. 1, it can be seen that when ε decreases, the exponential function $\theta(x, \varepsilon)$ approaches the unit step function very well. In Fig. 1, only the non-negative part of the unit step function and its approximation are shown. This is because the singular value of any matrix is non-negative, the definition of the matrix rank is only related to the non-negative part of the unit step function. Moreover, since $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$ can approximate the unit step function very well when x is non-negative, we propose to use this function for the rank approximation. In addition, there are other interesting candidates that are worthy of attention. One candidate is $1 - 1/(x + 1)$. In fact, this function has the same properties as $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$ [Prop. 2.1, [23]]. Therefore, our convergence result to be presented also holds for this function. Another candidate is the (offset and scaled) sigmoid function: $1/(1 + e^{-x})$. This function also approximates the unit step function well. However, since its gradient is more complicated to calculate compared to $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$, we prefer to use $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$ for the rank approximation. Thus, we find that $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$ has a twofold good property: i) the convergence of our method derived from this approximate function is provable and ii) it facilitates the calculation in the algorithm.

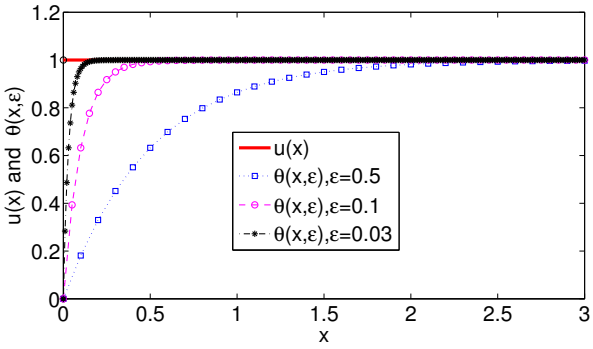


Fig. 1. Illustration of the approximated function of the unit step function. As the approximation parameter ε approaches zero, $\theta(x, \varepsilon)$ gets close to $u(x)$.

Then we define the approximation of the rank function associated with parameter ε as: $\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon) = \sum_{i=1}^n \theta(\sigma_i(\boldsymbol{W}), \varepsilon)$, and we establish the approximate problem of Problem (4) as follows:

$$\begin{aligned} \underset{\boldsymbol{W}\in\mathbb{S}^n}{\text{minimize}} \quad & \text{Tr}(\boldsymbol{A}_0\boldsymbol{W}) + \eta \cdot \widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon) \\ \text{subject to} \quad & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i\boldsymbol{W}) \le b_i,\ i = 1, \ldots, m. \end{aligned} \tag{5}$$

Denote the optimal solution of Problem (5) given $\varepsilon > 0$ as $\widetilde{\boldsymbol{W}}_{\varepsilon}^{opt}$. The relationship of the optimal solutions of Problem (4) and its approximate problem, Problem (5), is established in the following theorem.

Theorem 1. *For any decreasing sequence $\varepsilon_k \to 0$, if $\widetilde{\boldsymbol{W}}_{\varepsilon_k}^{opt} \to \widetilde{\boldsymbol{W}}$, $\widetilde{\boldsymbol{W}}$ is an optimal solution of Problem (4).*

Proof: See Appendix B. □

Theorem 1 shows that we can get the optimal solution of Problem (4) by first solving Problem (5) with a fixed parameter ε_k, whose initial value can be a large number. Then, we gradually decrease ε_k to make the rank approximation more accurate, and solve a sequence of Problem (5) for different parameters ε_k. Finally, the solution will converge to the optimal point of Problem (4). Though $\theta(x, \varepsilon)$ with a small ε is closer to the unit step function, its concavity is very large and may lead to computational difficulties in numerical calculations. This is because large concavity means that the slope of the curve decreases sharply as x increases and there may be a large number of local minimum points for the approximate problem. Thus, it is reasonable to solve Problem

(5) with decreasing values of ε_k to gradually approach the optimal solution of Problem (4).

Although Problem (5) is still nonconvex (i.e., minimization of a concave objective), it is a continuous and differentiable problem compared to Problem (4). In the following subsection, we will focus on solving Problem (5) by leveraging the MM technique. Note that it is hardly possible to achieve the optimal solution of Problem (5) due to its non-convexity. Thus, our target is to find a stationary point that is close to the optimal solution. Moreover, the performance of our method will be demonstrated by numerical results.

B. Majorization-minimization Method

The MM algorithm [25]–[27] (*a.k.a.* successive upper-bound minimization) tries to minimize $f(x)$ over $x \in \mathcal{X}$, where $\mathcal{X}$ is a closed convex set. This problem is difficult when the objective function $f(x)$ is nonconvex. The idea of the MM algorithm is to transform the original difficult problem into a series of simple problems, which is a generalization of many well known algorithms, such as the expectation maximization algorithm (EM) and the difference of convex functions (DC) algorithms [25]. In iteration $k+1$, the algorithm minimizes a surrogate function $g(x, x^k)$, in which x^k is the solution obtained from the kth iteration. This is to say that the following step is carried out to update the solution:

$$x^{k+1} = \arg\min_{x \in \mathcal{X}} g(x, x^k).$$

The MM algorithm solves the problem with surrogate objective functions iteratively until convergence. The difficulty of the MM algorithm lies in the construction of the surrogate functions, which is highly nontrivial and requires a careful design to facilitate the computation of each iteration. Moreover, in order to guarantee the convergence of the algorithm, the surrogate functions need to satisfy the following conditions as shown in [25].

C1 : $g(y, y) = f(y),\ \forall\ y \in \mathcal{X}$;
C2 : $g(x, y) \geq f(x),\ \forall\ x, y \in \mathcal{X}$;
C3 : $g'(x, y; d)|_{x=y} = f'(y; d),\ \forall\ d$ with $y + d \in \mathcal{X}$;
C4 : $g(x, y)$ is continuous in (x, y).

C1 and C2 require that the surrogate function is a tight upper bound of the original objective function. Define the directional derivative by

$$f'(x; d) = \liminf_{\lambda \to 0} \frac{f(x + \lambda d) - f(x)}{\lambda}.$$

Then C3 requires the same directional derivative of the original and surrogate functions at the point $x = y$.

C. MM-based Rank-one Solution Method

In Problem (5), the feasible set is a closed convex set and the objective function is composed of a convex function $\text{Tr}(\boldsymbol{A}_0\boldsymbol{W})$ and a nonconvex one $\eta \cdot \widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon)$. In order to solve Problem (5) by leveraging the MM algorithm, we continue to construct the surrogate function of $\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon)$ satisfying conditions C1–C4. More importantly, to ensure the computational efficiency after the substitution of the surrogate function, we tend to design the surrogate function as a linear function of $\boldsymbol{W}$ so that the resulting problem is an SDP problem in each iteration. Denote $\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon) = \sum_{i=1}^{n} \theta(\sigma_i(\boldsymbol{W}), \varepsilon) = h(\boldsymbol{\sigma}(\boldsymbol{W}))$ where $\boldsymbol{\sigma}(\boldsymbol{W}) = [\sigma_1(\boldsymbol{W}), \ldots, \sigma_n(\boldsymbol{W})]^T$, and the eigenvalue decomposition (EVD) of $\boldsymbol{W} \in \mathbb{S}_+^n$ is $\boldsymbol{P}\text{diag}(\boldsymbol{\sigma}(\boldsymbol{W}))\boldsymbol{P}^T$, in which the ith column of $\boldsymbol{P}$ is the eigenvector of $\boldsymbol{W}$. The choice of the surrogate functions is based on the following proposition.

Proposition 1: $\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon)$ is concave and its gradient at $\boldsymbol{W}$ is:

$$\frac{\partial \widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon)}{\partial \boldsymbol{W}} = \nabla \widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon) = \boldsymbol{P}\text{diag}(\boldsymbol{\beta})\boldsymbol{P}^T, \quad (6)$$

where $\boldsymbol{\beta} = \nabla h(\boldsymbol{\sigma}(\boldsymbol{W}))$ denotes the gradient of h at $\boldsymbol{\sigma}(\boldsymbol{W})$.

Proof: The proof follows Propositions 2 and 3 in [24]. □

Note that for a PSD matrix, the eigenvalues are equal to the singular values, which means $\sigma_i(\boldsymbol{W})$ is also the ith largest eigenvalue of $\boldsymbol{W} \in \mathbb{S}_+^n$. Proposition 1 shows the concavity of the approximate rank function $\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon)$ and its gradient. According to this, for some feasible matrix $\boldsymbol{W}_k$ obtained at the kth iteration, we have

$$\widetilde{\text{rank}}(\boldsymbol{W}, \varepsilon) \leq \widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon) + \langle \nabla \widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon), \boldsymbol{W} - \boldsymbol{W}_k \rangle, \quad (7)$$

where the gradient is given by the following equation:

$$\nabla \widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon) = \frac{1}{\varepsilon} \boldsymbol{P}\ \text{diag}(e^{-\sigma_1(\boldsymbol{W}_k)/\varepsilon}, \cdots, e^{-\sigma_n(\boldsymbol{W}_k)/\varepsilon}) \boldsymbol{P}^T,$$

for $\theta(x, \varepsilon) = 1 - e^{-x/\varepsilon}$. Consequently, we choose to use

$$g(\boldsymbol{W}, \boldsymbol{W}_k) = \text{Tr}(\boldsymbol{A}_0\boldsymbol{W}) + \eta \cdot \langle \nabla \widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon), \boldsymbol{W} \rangle$$

as the surrogate objective function for Problem (5) at a specific point $\boldsymbol{W}_k$, where the constant term $\eta(\widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon) - \langle \nabla \widetilde{\text{rank}}(\boldsymbol{W}_k, \varepsilon), \boldsymbol{W}_k \rangle)$ from the right-hand side of (7) is ignored because it will not affect the optimal solution.

As a result, if we denote the optimal solution obtained by the MM algorithm at the kth iteration as $\boldsymbol{W}_k^{opt}$, at the $k+1$th iteration, the following problem is solved in order to update the solution:

$$\begin{aligned} \underset{\boldsymbol{W} \in \mathbb{S}^n}{\text{minimize}} \quad & g(\boldsymbol{W}, \boldsymbol{W}_k^{opt}) \\ \text{subject to} \quad & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i\boldsymbol{W}) \leq b_i,\ i = 1, \ldots, m. \end{aligned} \quad (8)$$

Note that based on the surrogate function $g(\boldsymbol{W}, \boldsymbol{W}_k^{opt})$, Problem (8) is an SDP problem and can be solved efficiently.

Next we will explain how to get an initial point for our iterative algorithm. It is well known that the nuclear norm $||\boldsymbol{W}||_*$ is a convex approximation of the rank function and $||\boldsymbol{W}||_* = \text{Tr}(\boldsymbol{W})$ for $\boldsymbol{W} \in \mathbb{S}_+^n$. Thus, we solve the following problem to initialize our algorithm and denote its optimal solution as $\boldsymbol{W}_0^{opt}$,

$$\begin{aligned} \underset{\boldsymbol{W} \in \mathbb{S}^n}{\text{minimize}} \quad & \text{Tr}(\boldsymbol{A}_0\boldsymbol{W}) + \eta \cdot \text{Tr}(\boldsymbol{W}) \\ \text{subject to} \quad & \boldsymbol{W} \succeq \boldsymbol{0},\ \text{Tr}(\boldsymbol{A}_i\boldsymbol{W}) \leq b_i,\ i = 1, \ldots, m. \end{aligned} \quad (9)$$

Finally, our solution framework is summarized in Algorithm 1. For a given rank penalty coefficient η, there are two layers of loops in the algorithm. The inner-loop (Step 5 to Step 10) carries out the MM algorithm for a fixed approximation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

Algorithm 1 MM-based rank-one solution method for QCQPs

Input: $\eta > 0$, $\varepsilon_0 > 0$, $\alpha > 1$, $\delta_1 > 0$, $\delta_2 > 0$ are given

1: Solve Problem (9) to get $\boldsymbol{W}_0^{opt}$,
2: $l = 0$, $\varepsilon = \varepsilon_0$, $d_1 = \delta_1 + 1$,
3: **while** $d_1 > \delta_1$ **do**
4: $k = 0, \widetilde{\boldsymbol{W}}_0 = \boldsymbol{W}_l^{opt}$, $d_2 = \delta_2 + 1$,
5: **while** $d_2 > \delta_2$ **do**
6: Solve Problem (8) at point $\widetilde{\boldsymbol{W}}_k$ to get the optimal
7: solution $\widetilde{\boldsymbol{W}}_{k+1}$,
8: $d_2 = ||\widetilde{\boldsymbol{W}}_{k+1} - \widetilde{\boldsymbol{W}}_k||_F / ||\widetilde{\boldsymbol{W}}_k||_F$,
9: $k = k + 1$,
10: **end while**
11: $\boldsymbol{W}_{l+1}^{opt} = \widetilde{\boldsymbol{W}}_k$,
12: $d_1 = ||\boldsymbol{W}_{l+1}^{opt} - \boldsymbol{W}_l^{opt}||_F / ||\boldsymbol{W}_l^{opt}||_F$,
13: $l = l + 1$, $\varepsilon = \varepsilon / \alpha$.
14: **end while**
15: **if** rank$(\boldsymbol{W}_l^{opt}) > 1$ **then**
16: $\eta = 2\eta$, and go back to Step 1.
17: **end if**

Output: $\boldsymbol{W}_l^{opt}$

parameter ε and the outer-loop (Step 3 to Step 14) decreases ε at each iteration with a factor of α, in order to make the approximate function gradually approach the exact rank function. Based on our numerical tests, α can be chosen within the range $[2, 10]$. After the outer loop terminates, the algorithm checks the rank of the solution. If the rank is larger than one, the penalty coefficient will be increased (here we double it as an example) until a rank-one solution is output by the algorithm. When the penalty coefficient is large enough, this algorithm can guarantee to yield a rank-one solution eventually.

The stopping criteria for the two loops are based on d_1 and d_2, which are the relative distances of solutions obtained by successive iterations in the outer and inner loops, respectively. Moreover, δ_1 and δ_2 are two small numbers, which can be set to be 10^{-3} as an example.

The convergence of Algorithm 1 is ensured by the following theorem.

Theorem 2. *The solution obtained by Algorithm 1 is a stationary point of Problem (2).*

Proof: See Appendix C. □

IV. Global Search based on local smoothing

With Algorithm 1, we are able to find a stationary point, which is usually a local minimum point of Problem (2) if it is not a saddle point. Therefore, Algorithm 1 can be considered as a local search method for the SDP relaxation of QCQPs. In order to further improve the result, our method can be combined with any other global search algorithm. We propose to use the local smoothing technique [28] in this section.

A. Local smoothing transformation and its approximation

Consider a general optimization problem as:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad f(\boldsymbol{x}), \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{X}, \tag{10}$$

where $f(\boldsymbol{x})$ is nonconvex and has many local minimum points. Suppose that there is a local search algorithm $LS(\boldsymbol{x})$, which is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$, such that $\boldsymbol{y} = LS(\boldsymbol{x})$ is a local minimum point. Furthermore, if we denote the composed function $c(\boldsymbol{x}) = f(LS(\boldsymbol{x}))$, the following problem is equivalent to Problem (10):

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad c(\boldsymbol{x}), \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{X}. \tag{11}$$

In fact, many global optimization techniques (such as Multi-start and clustering method) reduce to the uniform random sampling applied to $c(\boldsymbol{x})$ [28]. Because the new objective function $c(\boldsymbol{x})$ is piecewise constant and hard to solve, a local smoothing transformation is applied to $c(\boldsymbol{x})$ such that the information of descent directions can be captured.

Given a Gaussian smoothing kernel $\psi(t) : \mathbb{R} \to \mathbb{R}$, defined as $\psi(t) = \exp(-\frac{t^2}{2\rho^2})$ where the multiplicative constant is ignored, the ψ-transform of $c(\boldsymbol{x})$ on a local region $B(\boldsymbol{x}_0, r) = \{\boldsymbol{x} : ||\boldsymbol{x} - \boldsymbol{x}_0|| \le r\}$ is

$$c_\psi^B(\boldsymbol{x}) = \frac{\int_{B(\boldsymbol{x}_0,r)} c(\boldsymbol{z})\psi(||\boldsymbol{z} - \boldsymbol{x}||)d\boldsymbol{z}}{\int_{B(\boldsymbol{x}_0,r)} \psi(||\boldsymbol{z} - \boldsymbol{x}||)d\boldsymbol{z}}. \tag{12}$$

The effect of this smoothing transformation is similar to a filter, which replaces the value of each point by the weighted average of its neighbors and the degree of smoothing is controlled by ρ. Leveraging the smooth function $c_\psi^B(\boldsymbol{x})$, the information of the descent direction of the original piecewise constant function can be obtained. Since the analytical expression of $c(\boldsymbol{z})$ is not available, an estimation of (12) is calculated based on K sample points $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K$, which are uniformly drawn from $B(\boldsymbol{x}_0, r)$:

$$\hat{c}_\psi^B(\boldsymbol{x}) = \frac{\sum_{i=1}^K c(\boldsymbol{z}_i)\psi(||\boldsymbol{z}_i - \boldsymbol{x}||)}{\sum_{i=1}^K \psi(||\boldsymbol{z}_i - \boldsymbol{x}||)}. \tag{13}$$

Note that both the one-dimensional and high-dimensional Gaussian kernel smoothing techniques have been widely used for signal and image smoothing [28]–[30]. Moreover, a n-dimensional Gaussian kernel smoothing can be done by applying one-dimensional smoothing n times in each direction [29]. Therefore, for the high-dimensional $\boldsymbol{x}$, the Gaussian kernel smoothing can still be applied to get a qualitative description of very complex functions.

B. Global search algorithm

For Problem (5), we have found a local minimum point $\boldsymbol{x}_0^{opt}$ where $\boldsymbol{x}_0^{opt}(\boldsymbol{x}_0^{opt})^T = \boldsymbol{W}_l^{opt}$, which is the output of Algorithm 1. In order to conduct a global exploration, a uniform random sampling is carried out in the neighborhood region of $\boldsymbol{x}_0^{opt}$: $B(\boldsymbol{x}_0^{opt}, r)$, and the sampling points are denoted as $\boldsymbol{x}_i, i = 1, \ldots, K$. Each point $\boldsymbol{x}_i$ is considered as a starting point and Algorithm 1 is applied to find a local minimum $\boldsymbol{z}_i$, with the corresponding objective value denoted as $c(\boldsymbol{z}_i)$. Then, the estimated local smoothing function is constructed by Eq. (13). Denote a local minimum of $\hat{c}_\psi^B(\boldsymbol{x})$ as $\boldsymbol{x}_1^{opt}$, which is considered as a new center and the procedure is repeated until the maximum number of iterations J is reached. The global search algorithm is summarized in Algorithm 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

Algorithm 2 Global search based on local smoothing

Input: r, J, K are given

Use Algorithm 1 to find a local minimum $\boldsymbol{x}_0^{opt}$ where $\boldsymbol{x}_0^{opt}(\boldsymbol{x}_0^{opt})^T = \boldsymbol{W}_l^{opt}$;
$j = 0$;
while $j < J$ **do**
  $i = 1$
  while $i \leq K$ **do**
    Randomly sample a point $\boldsymbol{x}_i$ in $B(\boldsymbol{x}_j^{opt}, r)$;
    Take $\boldsymbol{x}_i$ as a starting point and use Algorithm 1 to find a local minimum $\boldsymbol{z}_i$; $i = i + 1$;
  end while
  Use $\boldsymbol{z}_i, i{=}1, \ldots, K$ to construct $\hat{c}_\psi^B(\boldsymbol{x})$ based on (13);
  For $\hat{c}_\psi^B(\boldsymbol{x})$, find a local minimum $\boldsymbol{x}_{j+1}^{opt}$; $j = j + 1$;
end while

Output: $\boldsymbol{x}_J^{opt}$

V. Application to OPF problems and case studies

A. Classical OPF Problem formulation and SDP relaxation

OPF is a generic term that describes a broad class of problems in which we seek to optimize a specific objective function (e.g., generation cost or power loss) while satisfying network and operational constraints of power systems [31]. In this section, we will first introduce the conventional and the rank constrained formulations of OPF problems. Then, case studies are carried out for power grid testbeds.

Consider a power network of n buses. We denote the set of buses and the set of generator buses as $\mathcal{N} = \{1, 2, \ldots, n\}$ and $\mathcal{G} \subseteq \mathcal{N}$, respectively. Also, the set of transmission lines is denoted as $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. The π-model is used for transmission lines, in which the series impedance of the line (i, j) is z_{ij} and $y_{ij} = 1/z_{ij}$. The shunt admittance at both ends of the line (i, j) is denoted as $0.5\hat{y}_{ij}$ and the total admittance-to-ground at bus i is y_{ii}. We denote the admittance matrix of the network as $\boldsymbol{Y}$ whose (i, j)th element is $y_{ii} + \sum_{k\in\mathcal{N}(i)} y_{ik}$ if $i = j$ and $-y_{ij}$ otherwise. Then, the OPF problems can be formulated as follows.

$$\underset{\boldsymbol{P_G},\boldsymbol{Q_G},\boldsymbol{V}}{\text{minimize}} \quad \sum_{i\in\mathcal{G}} f_i(P_{G_i})$$

$$\text{s.t. } P_{G_i} - P_{D_i} = \text{Re}(y_{ii}^*|V_i|^2) + \sum_{j\in\mathcal{N}(i)} \text{Re}\{V_i(V_i^* - V_j^*)y_{ij}^*\}, \quad \forall i \in \mathcal{N}, \tag{14a}$$

$$Q_{G_i} - Q_{D_i} = \text{Im}(y_{ii}^*|V_i|^2) + \sum_{j\in\mathcal{N}(i)} \text{Im}\{V_i(V_i^* - V_j^*)y_{ij}^*\}, \quad \forall i \in \mathcal{N}, \tag{14b}$$

$$P_{G_i}^{min} \leq P_{G_i} \leq P_{G_i}^{max}, \quad \forall i \in \mathcal{G}, \tag{14c}$$

$$Q_{G_i}^{min} \leq Q_{G_i} \leq Q_{G_i}^{max}, \quad \forall i \in \mathcal{G}, \tag{14d}$$

$$V_i^{min} \leq |V_i| \leq V_i^{max}, \quad \forall i \in \mathcal{N}, \tag{14e}$$

$$|S_{ij}| \leq S_{ij}^{max}, \quad \forall (i, j) \in \mathcal{E}, \tag{14f}$$

where $P_{G_i} + jQ_{G_i}$ is the complex generation output at bus i and it is zero for $i \in \mathcal{N}\backslash\mathcal{G}$. $P_{D_i} + jQ_{D_i}$ is the load at bus i. Also, $S_{ij} = P_{ij} + jQ_{ij}$ is the complex power flow on line (i, j) and V_i is the complex voltage at bus i. The vectors of $\{P_{G_i}\}_{i\in\mathcal{G}}, \{Q_{G_i}\}_{i\in\mathcal{G}}$ and $\{V_i\}_{i\in\mathcal{N}}$ are denoted by $\boldsymbol{P_G}, \boldsymbol{Q_G}$ and $\boldsymbol{V}$, respectively. In addition, $\mathcal{N}(i)$ denotes the set of all buses that are directly connected to bus i. In Problem (14), the objective function is a summation of all the generation cost. As for constraints, Eq. (14a) and Eq. (14b) are power flow balance constraints. In addition, Eq. (14c)–Eq.(14f) require that the active power, reactive power, voltage magnitudes and apparent power flows are within the corresponding operating limits, respectively.

It can be noted that Problem (14) is indeed a QCQP and it can be reformulated as a rank constrained SDP problem [1] by defining

$$Y_i = e_i e_i^T \boldsymbol{Y}, \; Y_{ij} = (0.5\hat{y}_{ij} + y_{ij})e_i e_i^T - (y_{ij})e_i e_j^T,$$

$$\boldsymbol{Y}_i = \frac{1}{2}\begin{bmatrix} \text{Re}\{Y_i + Y_i^T\} & \text{Im}\{Y_i^T - Y_i\} \\ \text{Im}\{Y_i - Y_i^T\} & \text{Re}\{Y_i + Y_i^T\} \end{bmatrix},$$

$$\overline{\boldsymbol{Y}}_i = -\frac{1}{2}\begin{bmatrix} \text{Im}\{Y_i + Y_i^T\} & \text{Re}\{Y_i - Y_i^T\} \\ \text{Re}\{Y_i^T - Y_i\} & \text{Im}\{Y_i + Y_i^T\} \end{bmatrix},$$

$$\boldsymbol{Y}_{ij} = \frac{1}{2}\begin{bmatrix} \text{Re}\{Y_{ij} + Y_{ij}^T\} & \text{Im}\{Y_{ij}^T - Y_{ij}\} \\ \text{Im}\{Y_{ij} - Y_{ij}^T\} & \text{Re}\{Y_{ij} + Y_{ij}^T\} \end{bmatrix},$$

$$\overline{\boldsymbol{Y}}_{ij} = -\frac{1}{2}\begin{bmatrix} \text{Im}\{Y_{ij} + Y_{ij}^T\} & \text{Re}\{Y_{ij} - Y_{ij}^T\} \\ \text{Re}\{Y_{ij}^T - Y_{ij}\} & \text{Im}\{Y_{ij} + Y_{ij}^T\} \end{bmatrix},$$

$$M_i = \begin{bmatrix} e_i e_i^T & 0 \\ 0 & e_i e_i^T \end{bmatrix},$$

$$\boldsymbol{X} = [\text{Re}\{\boldsymbol{V}\}^T \; \text{Im}\{\boldsymbol{V}\}^T]^T, \; \boldsymbol{W} = \boldsymbol{X}\boldsymbol{X}^T.$$

With this, Problem (14) can be equivalently converted to:

$$\underset{\boldsymbol{P_G},\boldsymbol{Q_G},\boldsymbol{W}}{\text{minimize}} \quad \sum_{i\in\mathcal{G}} f_i(P_{G_i})$$

$$\text{s.t. } P_{G_i} - P_{D_i} = \text{Tr}(\boldsymbol{Y}_i\boldsymbol{W}), \quad \forall i \in \mathcal{N}, \tag{15a}$$

$$Q_{G_i} - Q_{D_i} = \text{Tr}(\overline{\boldsymbol{Y}}_i\boldsymbol{W}), \quad \forall i \in \mathcal{N}, \tag{15b}$$

$$P_{G_i}^{min} \leq P_{G_i} \leq P_{G_i}^{max}, \quad \forall i \in \mathcal{G}, \tag{15c}$$

$$Q_{G_i}^{min} \leq Q_{G_i} \leq Q_{G_i}^{max}, \quad \forall i \in \mathcal{G}, \tag{15d}$$

$$(V_i^{min})^2 \leq \text{Tr}(M_i\boldsymbol{W}) \leq (V_i^{max})^2, \quad \forall i \in \mathcal{N}, \tag{15e}$$

$$|\text{Tr}(\boldsymbol{Y}_{ij}\boldsymbol{W}) + j\text{Tr}(\overline{\boldsymbol{Y}}_{ij}\boldsymbol{W})| \leq S_{ij}^{max}, \; \forall (i, j) \in \mathcal{E}, \tag{15f}$$

$$\boldsymbol{W} \succeq 0, \tag{15g}$$

$$\text{rank}(\boldsymbol{W}) = 1. \tag{15h}$$

It can be easily shown that when each cost function $f_i(P_{G_i})$ is linear or quadratic in P_{G_i}, Problem (15) can be transformed to a standard SDP problem together with a rank-one constraint. Therefore, our algorithm can be applied to solve the OPF problems. In the following subsections, we will show the simulation results of applying our method to the OPF problems associated with different power system testbeds.

Since the general QCQPs are NP-hard (so are the OPF problems [1]), it is difficult to show whether the rank-one solution obtained by our method is globally optimal or not. However, we can measure the sub-optimality by comparing the objective value of the optimal solution obtained from our method (denoted as $\hat{f}^{opt}$) with the lower bound (denoted as LB) [13], which is obtained by solving the SDP relaxation problem (3), without the rank-one constraint. The sub-optimality degree is calculated by $\zeta = \frac{\hat{f}^{opt} - LB}{\hat{f}^{opt}} \times 100\%$. In

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

case $\hat{f}^{opt} = LB$, i.e., $\zeta = 0$, it means that the solution is globally optimal and the relaxation is exact. It has been shown that for OPF problems, the SDP relaxation is exact if and only if the duality gap is zero [1].

B. Zero duality gap cases

Nine power system testbeds are considered in this subsection. The first three are small systems with three or four buses (Systems 1, 2, and 3 in [1]), for which the objective function is the power loss. The detailed specifications and constraints for these systems can be found in Tables I and II in [1], respectively. Another two radial systems are also considered. One is 32-bus [32] and another is 34-bus [33]. For both systems, the objective is to minimize the power loss and there is no limit on the apparent power flow. In addition, we have conducted simulations on IEEE systems with 14, 30, 57 and 118 buses, which are archived at [34] and the data are extracted from the software toolbox MATPOWER [35]. In these IEEE systems, the objective functions are the total generation costs, which are quadratic functions in the amount of active power generation. The simulation results for these nine systems are shown in the first nine rows in Table I, where we denote the objective value obtained from MATPOWER as f^{MP} and the second largest eigenvalues of the final solutions as SLE, respectively. For Algorithm 1, the parameters are set as follows: $\delta_1 = 10^{-4}$, $\delta_2 = 10^{-4}$ and $\alpha = 2$. As for the initial value of η, it is set as 1, 1, 1, 1, 1, 20, 20, 20 and 320 for the nine systems, respectively. Note that we set a large initial η for some systems just for reducing the number of iterations until a *large enough* η can be reached.

From Table I, it can be seen that for all these nine cases, our method can always yield a rank-one solution (thus feasible), and the performance of our method is as good as that of MATPOWER. Additionally, all of them achieve the global optimal solution except for the IEEE 118-bus system, where the sub-optimality degree is extremely small. This small sub-optimality degree for the IEEE 118-bus system is also observed in [2]. Note that even when the duality gap is zero, there may exist a large number of high rank infeasible solutions [1], while our method can find the rank-one feasible solution without suffering any optimality loss. Moreover, for all these cases, we are able to find the global optimal point by Algorithm 1 and thus, there is no need to apply Algorithm 2 here.

C. Non-zero duality gap cases

In the previous part, we have shown the practicability of our method for power systems with zero duality gap. Although the OPF problems for a significant number of practical systems have been tested to have zero duality gap, the authors in [15] explored a counterexample: the duality gap could be non-zero in case of strict line-flow constraints. In this regard, additional simulations are conducted.

Consider a three-bus system whose topology and load demands are shown in Fig. 2. Each bus has active and reactive power demands with the base of 100 MVA. Also, Buses 1 and 2 both have generators which can produce active and reactive power, while a synchronous condenser which can only produce reactive power is connected to Bus 3. It is assumed that there is no power limit for all the condenser and generators. For Buses 1 and 2, the generation cost is quadratic: $f_i(P_{G_i}) = c_{i2}P_{G_i}^2 + c_{i1}P_{G_i} + c_{i0}$, where the coefficients c_{i2}, c_{i1}, and c_{i0} are \$0.11 per $(\text{MWh})^2$, \$5 per MWh and \$0, respectively for Bus 1 and \$0.085 per $(\text{MWh})^2$, \$1.2 per MWh and \$0, respectively for Bus 2. Bus 3 has no generation cost and the allowable range for voltage magnitudes at all the buses is [0.9, 1.1] in p.u. As for the parameters in p.u. in the π-model of the transmission lines, $z_{13} = 0.065 + j0.620$, $z_{23} = 0.025 + j0.750$, and $z_{12} = 0.042 + j0.900$ are the series impedances. What's more, $\hat{y}_{13} = j0.450$, $\hat{y}_{23} = j0.700$ and $\hat{y}_{12} = j0.300$ are the total shunt susceptances for each line. Then, we will impose constraints on the line-flow limits. According to the different apparent power flow limits on both ends of the transmission lines, we have generated three different cases denoted as Cases 1, 2 and 3 shown in Table II, where the transmission line connecting Bus i and Bus j is denoted as l_{ij}. In addition, we have conducted a simulation for a larger system: the New England 39-bus [34] and compare the performance of our algorithm with MATPOWER in Table II. The parameter setting in Algorithm 2 is as follows: $J = 10$, $K = 1000$ and $r \in \{0.1, 0.2, 0.3\}$.

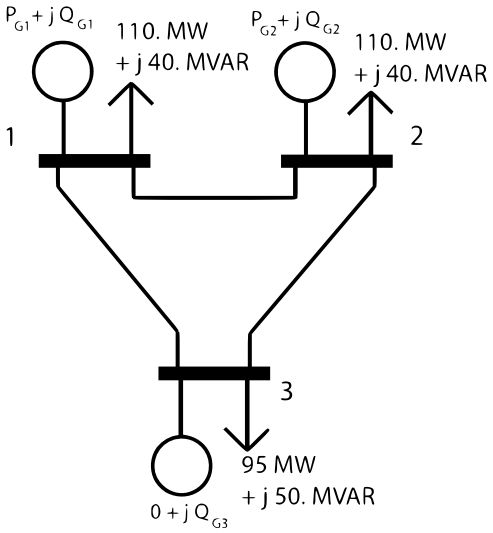


Fig. 2. A three-bus system [15].

TABLE II
CASES WITH STRICT LINE-FLOW LIMITS.

	Line-flow limits (MVA)		
	l_{12}	l_{13}	l_{23}
Case 1	$+\infty$	$+\infty$	50
Case 2	$+\infty$	$+\infty$	45
Case 3	25	$+\infty$	50

The simulation results for these four cases are shown in the last four rows of Table I. Furthermore, the changes of the second largest eigenvalues (SLE) of the solutions with increasing η are shown in Fig. 3. In Fig. 3, it can be seen that with the increase of η, the second largest eigenvalues decreases. Also, when η becomes large enough, the SLE is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

TABLE I
SIMULATION RESULTS OF OPF PROBLEMS.

	LB	$\hat{f}^{opt}$	f^{MP}	ζ	SLE	rank	η	n_ε	n_η	Run. time (sec)
System 1 [1]	0.2259	0.2259	0.2259	0	3.4E-10	1	1	6	1	5.4
System 2 [1]	0.1588	0.1588	0.1588	0	9.9E-10	1	1	7	1	4.4
System 3 [1]	0.3877	0.3877	0.3877	0	6.2E-10	1	1	7	1	5.0
32-bus system [32]	202.68	202.68	202.68	0	1.3E-9	1	1	9	1	226.8
34-bus system [33]	12.92	12.92	12.92	0	7.5E-9	1	1	9	1	228.5
IEEE-14	8081.53	8081.53	8081.53	0	4.5E-11	1	20	8	1	22.1
IEEE-30	576.89	576.89	576.89	0	2.8E-6	1	20	10	1	192.3
IEEE-57	41737.8	41737.8	41737.8	0	3.3E-5	1	20	10	1	3015
IEEE-118	129654.4	129660.7	129660.7	0.0046%	6.3E-7	1	320	11	1	436331
Case 1	5789.9	5812.6	5812.6	0.39%	1.4E-12	1	320	6	4	36.4
Case 2	5869.9	6038.3	6038.3	2.79%	7.5E-14	1	320	9	4	46.4
Case 3	5793.6	5831.4	5831.4	0.65%	8.4E-10	1	640	8	5	60.5
New England 39-bus	41862.1	41864.2	41864.2	0.005%	3.8E-7	1	320	10	4	1716

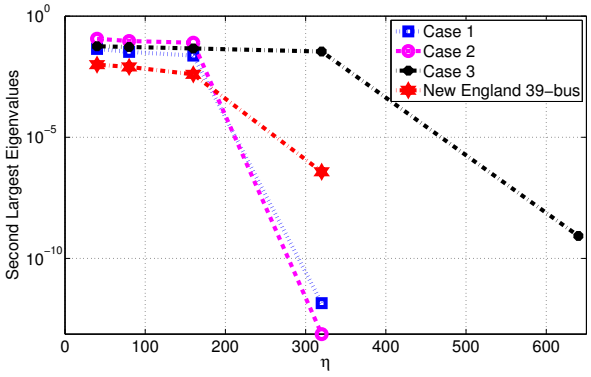


Fig. 3. Second largest eigenvalues for different η.

very small such that the solutions are rank-one. From Table I, we observe that the sub-optimality degrees are all very small and more importantly, for all the cases, our method is successful to find rank-one solutions which are as good as the performance of MATPOWER. Note that compared with the penalization method in [2], our method is more general and there is no need to fine tune any specific parameter in contrast to that in [2]. Therefore, we have shown that the performance of our method is good for these non-zero duality gap cases. In addition, for these four cases, we observe that the global search algorithm yields the same result as that of Algorithm 1. This means that for the cases we have tested, the solutions are not further improved by Algorithm 2, which can be considered as a support for the good performance of Algorithm 1.

Note that although SDP is convex, its computational complexity is still high for large systems. Therefore, the largest systems considered in our experiments are the IEEE 118-bus system and the New England 39-bus system for zero and non-zero duality gap cases, respectively. If larger systems are considered, a distributed algorithm for solving the large scale SDP problems should be studied such as in [36].

D. Computation information

The problems in the case studies are solved by the CVX solver on a single machine with an i5 dual-core processor. The computation information are shown for the OPF problems in the last three columns of Table I. Specifically, we show the total running time for the rank-one solutions to be obtained as well as two numbers of iterations: i) the number of iterations for increasing η, which is denoted as n_η and ii) the number of iterations for decreasing ε with a fixed η, which is denoted as n_ε. From the table, it can be seen that the running time varies from several seconds to tens of thousands of seconds. While for the interior point solver in MATPOWER, it takes less than one second for all of these cases. Therefore, we admit that before any distributed algorithm is designed for our rank-constrained SDP algorithm, the computational complexity is higher than the interior point solver in MATPOWER.

We have to emphasize that although in this section we have shown the different cases of applying our method to OPF problems, we never state that our method is superior to any state-of-the-art OPF solvers. In contrast, we want to underline the generality of our solution framework. That is, whenever a problem is formulated as a QCQP, our method can be readily applied to get a practically meaningful solution, which has a great chance to be at least locally optimal. As a result, the solution obtained by our method can be regarded as a good starting point for more sophisticated algorithms which can explore the specific structures of different problems.

E. Other update rules for the penalty parameter

In Algorithm 1, the penalty parameter η is doubled at each iteration and from Table I, it can be seen that the "large enough" values of η for Cases 1, 2 and 3 are equal to 320, 320 and 640, respectively. In comparison, we have tried another update rule that is *linear*. Specifically, the initial values of η are also equal to 20, but at each step, we increase η by a constant step size of 20. After the rank-one solutions are obtained, we find that the η values are equal to 220, 240 and 380, respectively for the three cases. It can be seen that the linear update rule has found smaller η values. This is normal because the doubling update rule is more aggressive to find the large η. However, an important finding is that both rules can obtain the rank-one solution with the *same* objective value. Therefore, we can know that as long as the penalty parameter is large enough, the solution is not very sensitive to the value of η. According to this, the doubling update rule can be a good way to get a rank-one solution with less iterations.

We have shown the successful application of our method

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

to the OPF problems. In the next section, we will show the application of our method to the state estimation problems.

VI. Application to State Estimation Problems and Case Studies

First, we will briefly introduce how our method can be used to solve the state estimation problem and then show the numerical results on the IEEE 14-bus system.

The goal of state estimation (SE) problems is to compute the complex bus voltages from a set of redundant measurements such as the active power injection, reactive power flow and voltage magnitude, etc [37], [38]. To formulate the SE problem, we follow the same notation in the example of OPF problems. Consider a power network with n buses. Denote the complex voltage at bus i as V_i and the vector of $\{V_i\}_{i=1,\cdots,n}$ as $\boldsymbol{V}$. A subset of the following system variables are measured:

1) The active and reactive power injections at bus i: P_i^{inj} and Q_i^{inj};
2) The active and reactive power flows at both ends of line (i,j): P_{ij}, P_{ji}, Q_{ij} and Q_{ji};
3) The voltage magnitude at bus i: $|V_i|$.

Suppose that there are M number of measurements denoted as $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \cdots, \kappa_M,)$. Then, the m-th entry of $\boldsymbol{\kappa}$ can be modeled as [39]:

$$\kappa_m = h_m(\boldsymbol{V}) + \epsilon_m, \tag{16}$$

where h_m denotes the quadratic dependence of $\kappa_m(\cdot)$ on $\boldsymbol{V}$. Each ϵ_m is assumed to be an independent Gaussian random variable with zero mean: i.e., $\epsilon_m \sim N(0, \nu_m^2)$, where ν_m denotes the corresponding standard deviation. Then, the SE problem aims to find an estimate $\hat{\boldsymbol{V}}$ of $\boldsymbol{V}$ that best matches the measurement set $\boldsymbol{\kappa}$ according to the relationships in (16). The general formulation for SE problem is as follows:

$$\min_{\boldsymbol{V}} \quad J(\boldsymbol{V}) = \left(\sum_{m=1}^{M} \left| \frac{\kappa_m - h_m(\boldsymbol{V})}{\nu_m} \right|^p \right)^{1/p}, \tag{17}$$

where the different choices of p can result in different estimation criteria. For example, in this paper we choose $p = 2$ and this is the Weighted Least Square (WLS) estimation:

$$\min_{\boldsymbol{V}} \quad J(\boldsymbol{V}) = \sum_{m=1}^{M} \frac{1}{\nu_m^2} \Big(\kappa_m - h_m(\boldsymbol{V}) \Big)^2. \tag{18}$$

From the OPF example, we know that $h_m(\boldsymbol{V})$ is a quadratic function of $\boldsymbol{V}$ and by defining $\boldsymbol{X}$, $\boldsymbol{W}$, $\boldsymbol{Y}_i$, $\overline{\boldsymbol{Y}}_i$, $\boldsymbol{Y}_{ij}$, $\overline{\boldsymbol{Y}}_{ij}$ and M_i as in the OPF problem, all the measurements can be modeled by a linear relationship with the matrix variable $\boldsymbol{W}$:

$$\kappa_m = \text{Tr}(\boldsymbol{H}_m \boldsymbol{W}) + \epsilon_m,$$

where $\boldsymbol{H}_m$ denotes the corresponding measurement matrix for each κ_m. Therefore, the SE problem can be converted to a problem with $\boldsymbol{W}$ as the decision variable:

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & J(\boldsymbol{W}) = \sum_{m=1}^{M} \frac{1}{\nu_m^2} \Big(\kappa_m - \text{Tr}(\boldsymbol{H}_m \boldsymbol{W}) \Big)^2 \\ \text{subject to} \quad & \boldsymbol{W} \succeq 0, \ \text{rank}(\boldsymbol{W}) = 1. \end{aligned} \tag{19}$$

Based on the Schur's complement lemma [39], this problem can be further converted to a standard SDP form with a rank-one constraint:

$$\begin{aligned} \min_{\boldsymbol{W}, \boldsymbol{\tau}} \quad & \boldsymbol{\mu}^T \boldsymbol{\tau} \\ \text{s.t.} \quad & \begin{bmatrix} \tau_m & \kappa_m - \text{Tr}(\boldsymbol{H}_m \boldsymbol{W}) \\ \kappa_m - \text{Tr}(\boldsymbol{H}_m \boldsymbol{W}) & 1 \end{bmatrix} \succeq 0, m = 1, ..., M \\ & \text{rank}(\boldsymbol{W}) = 1. \end{aligned} \tag{20}$$

where $\boldsymbol{\mu} = [1/\nu_1^2, \cdots, 1/\nu_M^2]^T$ and $\boldsymbol{\tau} = [\tau_1, \cdots, \tau_M]^T$. Therefore, our algorithm can be applied to solve SE problems. Next, we will show the simulation results for the SE problem on the IEEE 14-bus system.

The measurements available are listed in Table III. The data of these measurements are generated in the following way. First, the OPF problem of this system is solved by MATPOWER (or our proposed method) to get the true values of these variables without noise. Then, based on the variance ν_m^2 for each measurement, we add some randomly sampled Gaussian noise to the true values to get the vector $\boldsymbol{\kappa}$. We have generated 10 different samples of the vector $\boldsymbol{\kappa}$ to see the performance of our method and the results are shown in Table IV. From the table, we can see that for each measurement data, our method is able to find the rank-one solutions with very small sub-optimality degrees. Therefore, we have shown the performance of applying our general solution framework to different QCQP problems in power systems and find that it can always obtain rank-one solutions.

VII. Conclusions

This paper has proposed a solution framework for the QCQPs in power systems. Different from existing works that apply SDP relaxation to QCQPs, our proposed method has been guaranteed to obtain a rank-one feasible solution. Moreover, it has been shown that our proposed algorithm can converge to a stationary point with a small gap of optimality. In addition, a global search algorithm based on local smoothing technique has been introduced to further improve the quality of solutions. In the case studies, we have successfully applied our method to OPF problems and SE problems, in which our proposed solution framework has achieved good performance based on the extensive numerical results and find that rank-one solutions can always be obtained.

TABLE III
SE measurements of IEEE 14-bus system.

	Bus index of measurements	ν_m
P_i^{inj}	1,2,3,4,5	0.015
Q_i^{inj}	1,2	0.015
$\lvert V_i \rvert$	2,3,6,8,10,14	0.01
	Line index of measurements	
P_{ij}	(1,2),(2,3),(4,7),(4,9),(5,6),(6,13),(7,9),(9,10),(9,14),(12,13),(4,2),(5,2),(5,4),(11,6)	0.02
Q_{ij}	(1,2),(2,3),(4,7),(4,9),(5,6),(6,13),(7,9),(12,13),(4,2),(5,2),(5,4),(11,6)	0.02

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2729082, IEEE Transactions on Smart Grid

TABLE IV
SE RESULTS OF IEEE 14-BUS SYSTEM.

Sample index	1	2	3	4	5
LB	15.30	3.34	17.12	5.22	10.50
$\hat{f}^{opt}$	15.42	3.38	17.73	5.42	11.01
ζ	0.75%	1.11%	3.46%	3.52%	4.65%
SLE	1.9E-9	1.1E-8	6.3E-9	4.1E-9	6.6E-10
η	640	640	640	640	640
Sample index	6	7	8	9	10
LB	15.54	5.50	5.67	1.46	1.71
$\hat{f}^{opt}$	16.46	5.85	6.11	1.57	1.86
ζ	5.59%	6.20%	7.23%	7.5%	7.69%
SLE	6.4E-9	3.1E-9	2.8E-9	4.6E-9	2.9E-9
η	640	640	640	640	640

APPENDIX A
INHOMOGENEOUS QCQPS

The standard form of general inhomogeneous QCQPs is:

$$\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} & \boldsymbol{x}^T\boldsymbol{A}_0\boldsymbol{x}+2\boldsymbol{a}_0^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x}+2\boldsymbol{a}_i^T\boldsymbol{x}\le b_i,\ i=1,\ldots,m. \end{array} \tag{21}$$

According to [12], Problem (21) can be homogenized as:

$$\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}}{\text{minimize}} & [\boldsymbol{x}^T\ t]\begin{bmatrix}\boldsymbol{A}_0 & \boldsymbol{a}_0\\ \boldsymbol{a}_0^T & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ t\end{bmatrix} \\ \text{subject to} & [\boldsymbol{x}^T\ t]\begin{bmatrix}\boldsymbol{A}_i & \boldsymbol{a}_i\\ \boldsymbol{a}_i^T & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ t\end{bmatrix}\le b_i,\ i=1,\ldots,m, \\ & t^2=1. \end{array}$$

Therefore, the SDP relaxation technique is applicable to general inhomogeneous QCQPs by homogenization.

APPENDIX B
PROOF OF THEOREM 1

The proof of this theorem is mainly from the Proposition 2.4 (c) in [23] and here we briefly introduce the key steps. First, it can be easily shown that the approximate function $\theta(x,\varepsilon)=1-e^{-x/\varepsilon}$ satisfies $\theta(x,\varepsilon)\le u(x)$ and $\lim\limits_{\varepsilon\to 0}\theta(x,\varepsilon)=u(x),\ \forall x\ge 0$. Therefore, we have $\widetilde{\text{rank}}(\boldsymbol{W},\varepsilon)\le \text{rank}(\boldsymbol{W})$ and $\lim\limits_{\varepsilon\to 0}\widetilde{\text{rank}}(\boldsymbol{W},\varepsilon)=\text{rank}(\boldsymbol{W})$ since $\widetilde{\text{rank}}(\boldsymbol{W},\varepsilon)=\sum_{i=1}^n\theta(\sigma_i(\boldsymbol{W}),\varepsilon)\le\sum_{i=1}^n u(\sigma_i(\boldsymbol{W}))=\text{rank}(\boldsymbol{W})$. Then let $\boldsymbol{W}^{opt}$ be an optimal solution of Problem (4), then, for each k, it holds that

$$\begin{aligned} &\text{Tr}(\boldsymbol{A}_0\boldsymbol{W}^{opt})+\eta\cdot\text{rank}(\boldsymbol{W}^{opt})\\ \ge\ &\text{Tr}(\boldsymbol{A}_0\boldsymbol{W}^{opt})+\eta\cdot\widetilde{\text{rank}}(\boldsymbol{W}^{opt},\varepsilon_k)\\ \ge\ &\text{Tr}(\boldsymbol{A}_0\widetilde{\boldsymbol{W}}_{\varepsilon_k}^{opt})+\eta\cdot\widetilde{\text{rank}}(\widetilde{\boldsymbol{W}}_{\varepsilon_k}^{opt},\varepsilon_k)\\ \overset{k\to\infty}{\longrightarrow}\ &\text{Tr}(\boldsymbol{A}_0\widetilde{\boldsymbol{W}})+\eta\cdot\text{rank}(\widetilde{\boldsymbol{W}})\\ \ge\ &\text{Tr}(\boldsymbol{A}_0\boldsymbol{W}^{opt})+\eta\cdot\text{rank}(\boldsymbol{W}^{opt}), \end{aligned}$$

where the last inequality comes from the feasibility of $\widetilde{\boldsymbol{W}}$ and the optimality of $\boldsymbol{W}^{opt}$, so we get the result.

APPENDIX C
PROOF OF THEOREM 2

Our algorithm is based on the MM technique and in this technique, the limit point of the iterates has been shown to be a stationary point when the surrogate functions satisfy conditions C1–C4 [25]. In Problem (5), the original objective is $\text{Tr}(\boldsymbol{A}_0\boldsymbol{W})+\eta\cdot\widetilde{\text{rank}}(\boldsymbol{W},\varepsilon)$, where the first part is convex while the second part is concave based on Proposition 1. Then from Eq. (7), we use the first order derivative of the concave part to construct the surrogate function $\text{Tr}(\boldsymbol{A}_0\boldsymbol{W})+\eta\cdot\left(\widetilde{\text{rank}}(\boldsymbol{W}_k,\varepsilon)+\langle\nabla\widetilde{\text{rank}}(\boldsymbol{W}_k,\varepsilon),\boldsymbol{W}-\boldsymbol{W}_k\rangle\right)$ for some feasible point $\boldsymbol{W}_k$. According to this, it can be easily verified that the surrogate functions satisfy conditions C1–C4, which guarantees that a stationary point of Problem (5) can be found by Algorithm 1. Based on Theorem 1 and that the rank of the solution is one, this solution is also a stationary point of Problem (2), which is our original problem. Consequently, we can get the result in Theorem 2

REFERENCES

[1] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, 2012.

[2] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1297–1307, 2016.

[3] E. Dall'Anese, S. V. Dhople, and G. B. Giannakis, "Photovoltaic inverter controllers seeking ac optimal power flow solutions," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2809–2823, 2016.

[4] M. Ghamkhari, A. Sadeghi-Mobarakeh, and H. Mohsenian-Rad, "Strategic bidding for producers in nodal electricity markets: A convex relaxation approach," *IEEE Trans. Power Syst.*, 2016.

[5] S. J. Kazempour, A. J. Conejo, and C. Ruiz, "Strategic generation investment using a complementarity approach," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 940–948, 2011.

[6] A. Nemirovski, C. Roos, and T. Terlaky, "On maximization of quadratic form over intersection of ellipsoids with common center," *Mathematical Programming*, vol. 86, no. 3, pp. 463–473, 1999.

[7] C. Chen, A. Atamtürk, and S. S. Oren, "A spatial branch-and-cut method for nonconvex qcqp with bounded complex variables," *Mathematical Programming*, pp. 1–29.

[8] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "The qc relaxation: A theoretical and computational study on optimal power flow," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3008–3018, 2016.

[9] H. Hijazi, C. Coffrin, and P. Van Hentenryck, "Polynomial sdp cuts for optimal power flow," in *Power Systems Computation Conference (PSCC), 2016.* IEEE, 2016, pp. 1–7.

[10] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM (JACM)*, vol. 42, no. 6, pp. 1115–1145, 1995.

[11] Z.-Q. Luo and T.-H. Chang, "4 sdp relaxation of homogeneous quadratic optimization: approximation," *Convex Optimization in Signal Processing and Communications*, p. 117, 2010.

[12] Z.-Q. Luo, W.-k. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, p. 20, 2010.

[13] S. H. Low, "Convex relaxation of optimal power flow, part ii: Exactness," *arXiv preprint arXiv:1405.0814*, 2014.

[14] A. Lam, A. Dominguez-Garcia, B. Zhang, and D. Tse, "Optimal distributed voltage regulation in power distribution networks," Tech. Rep., 2012.

[15] B. C. Lesieutre, D. K. Molzahn, A. R. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *Proc. 49th Ann. Allerton Conf. Communi., Control, Comput.*, 2011, pp. 1492–1499.

[16] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 199–211, 2015.

[17] D. K. Molzahn, C. Josz, I. A. Hiskens, and P. Panciatici, "A Laplacian-based approach for finding near globally optimal solutions to OPF problems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 305–315, 2017.

[18] D. Molzahn, C. Josz, I. Hiskens, and P. Panciatici, "Solution of optimal power flow problems using moment relaxations augmented with objective function penalization," in *Proc. IEEE 54th Annu. Conf. Decision Control*, 2015, pp. 31–38.

[19] S. You and Q. Peng, "A non-convex alternating direction method of multipliers heuristic for optimal power flow," in *Proc. of IEEE SmartGridComm.*, 2014, pp. 788–793.

[20] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[21] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[22] A. R. Conn, "Constrained optimization using a nondifferentiable penalty function," *SIAM Journal on Numerical Analysis*, vol. 10, no. 4, pp. 760–784, 1973.

[23] S. Bi, L. Han, and S. Pan, "Approximation of rank function and its application to the nearest low-rank correlation matrix," *Journal of Global Optimization*, vol. 57, no. 4, pp. 1113–1137, 2013.

[24] M. Malek-Mohammadi, M. Babaie-Zadeh, and M. Skoglund, "Iterative concave rank approximation for recovering low-rank matrices," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5213–5226, 2014.

[25] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[26] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *Journal of computational and graphical statistics*, vol. 9, no. 1, pp. 1–20, 2000.

[27] D. R. Hunter and K. Lange, "A tutorial on mm algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[28] B. Addis, M. Locatelli, and F. Schoen, "Local optima smoothing for global optimization," *Optimization Methods and Software*, vol. 20, no. 4-5, pp. 417–437, 2005.

[29] M. K. Chung, "Gaussian kernel smoothing," *Lecture notes*, pp. 1–10, 2012.

[30] A. Witkin, "Scale-space filtering: A new approach to multi-scale description," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, vol. 9, 1984, pp. 150–153.

[31] J. A. Momoh, R. Adapa, and M. E. El-Hawary, "A review of selected optimal power flow literature to 1993. I. nonlinear and quadratic programming approaches," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 96–104, 1999.

[32] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401–1407, 1989.

[33] T. Liu, X. Tan, B. Sun, Y. Wu, X. Guan, and D. H. K. Tsang, "Energy management of cooperative microgrids with p2p energy sharing in distribution networks," in *Proc. of IEEE SmartGridComm.*, 2015, pp. 410–415.

[34] University of Washington, Power Systems Test Case Archive. [Online]. Available: http://www2.ee.washington.edu/research/pstca/

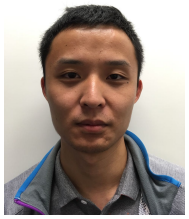
[35] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, 2011.

[36] R. A. Jabr, "Exploiting sparsity in sdp relaxations of the opf problem," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1138–1139, 2012.

[37] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Semidefinite programming for power system state estimation," in *Power and Energy Society General Meeting, 2012 IEEE*. IEEE, 2012, pp. 1–8.

[38] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

[39] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1039–1050, 2014.



Tian Liu (S'15) received the B.E. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2013, and the Engineering Diploma from Ecole Centrale Marseille, Marseille, France, in 2015. He is currently pursing the Ph.D. degree in electronic and computer engineering at the Hong Kong University of Science and Technology. His research interests are on optimization with applications in power systems.

Bo Sun (S'14) received the B.E. degree in electronic and information engineering from Harbin Institute of Technology, Harbin, China, in 2013. He is currently pursing the Ph.D. degree in electronic and computer engineering at the Hong Kong University of Science and Technology. His research interests are on queueing theory, Markov decision process and discrete convex analysis with applications in energy systems.

Danny H.K. Tsang (M'82-SM'00-F'12) received the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1989. Since 1992, he has been with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, where he is currently a Professor. His current research interests include Internet quality of service, peer-to-peer (P2P) video streaming, cloud computing, cognitive radio networks, and smart grids. Dr. Tsang served as a Guest Editor for the IEEE Journal of Selected Areas in Communications Special Issue on Advances in P2P Streaming Systems, an Associate Editor for the Journal of Optical Networking published by the Optical Society of America, and a Guest Editor of the IEEE Systems Journal. He currently serves as Technical Editor for the IEEE Communications Magazine.